The Journal of Chemical Physics

A simplified charge projection scheme for long-range electrostatics in *ab initio* QM/MM calculations

Cite as: J. Chem. Phys. **154**, 024115 (2021); https://doi.org/10.1063/5.0038120
Submitted: 19 November 2020 . Accepted: 17 December 2020 . Published Online: 13 January 2021

Xiaoliang Pan, Kwangho Nam, Evgeny Epifanovsky, Andrew C. Simmonett, Edina Rosta, and Yihan Shao

AIP Publishing

A simplified charge projection scheme for long-range electrostatics in *ab initio* QM/MM calculations

Cite as: J. Chem. Phys. 154, 024115 (2021); doi: 10.1063/5.0038120
Submitted: 19 November 2020 • Accepted: 17 December 2020 •
Published Online: 13 January 2021

View Online Export Citation CrossMark

Xiaoliang Pan,[1] Kwangho Nam,[2] Evgeny Epifanovsky,[3] Andrew C. Simmonett,[4] Edina Rosta,[5] and Yihan Shao[1,a)]

AFFILIATIONS

[1] Department of Chemistry and Biochemistry, University of Oklahoma, 101 Stephenson Pkwy, Norman, Oklahoma 73019, USA
[2] Department of Chemistry and Biochemistry, University of Texas at Arlington, Arlington, Texas 76019, USA
[3] Q-Chem, Inc., 6601 Owens Drive, Suite 105, Pleasanton, California 94588, USA
[4] National Institutes of Health–National Heart, Lung and Blood Institute, Laboratory of Computational Biology, Bethesda, Maryland 20892, USA
[5] Department of Physics and Astronomy, University College London, London WC1E 6BT, United Kingdom

[a)] Author to whom correspondence should be addressed: yihan.shao@ou.edu

ABSTRACT

In a previous work [Pan *et al.,* Molecules **23**, 2500 (2018)], a charge projection scheme was reported, where outer molecular mechanical (MM) charges [>10 Å from the quantum mechanical (QM) region] were projected onto the electrostatic potential (ESP) grid of the QM region to accurately and efficiently capture long-range electrostatics in *ab initio* QM/MM calculations. Here, a further simplification to the model is proposed, where the outer MM charges are projected onto inner MM atom positions (instead of ESP grid positions). This enables a representation of the long-range MM electrostatic potential via *augmentary charges* (AC) on inner MM atoms. Combined with the long-range electrostatic correction function from Cisneros *et al.* [J. Chem. Phys. **143**, 044103 (2015)] to smoothly switch between inner and outer MM regions, this new QM/MM-AC electrostatic model yields accurate and continuous *ab initio* QM/MM electrostatic energies with a 10 Å cutoff between inner and outer MM regions. This model enables efficient QM/MM cluster calculations with a large number of MM atoms as well as QM/MM calculations with periodic boundary conditions.

I. INTRODUCTION

Combined quantum mechanical and molecular mechanical (QM/MM) calculations have been widely used in the study of molecular solvation, ligand–receptor binding, chemical/enzyme reactions, photochemistry, and photobiology.[1–11] Within QM/MM calculations, a central region of interest in a system (such as a reactive site and a binding pocket) containing up to a few hundred atoms is designated as the quantum mechanical (QM) region, whose internal nuclear and electronic motions are subjected to rigorous QM modeling. Meanwhile, the remaining atoms of the system constitute the molecular mechanical (MM) region, whose internal (nuclear) motions are captured by classical MM force fields. Finally, the interactions between QM and MM regions consist of three types of terms: QM/MM covalent bonding, QM/MM electrostatics, and QM/MM van der Waals (vdW) interactions.

Out of the three types of interactions mentioned above, QM/MM electrostatics is our primary concern in this work. QM/MM covalent and vdW interactions are significant only for QM-MM atom pairs in close contact. Nevertheless, it should be noted that link-atom, double link atom,[12] local SCF,[13,14] generalized hybrid orbitals,[15,16] pseudobond,[17,18] frozen orbitals,[19,20] Yin-Yang atom,[21] and other approaches were developed to handle covalent bonds across the QM/MM interface. This is usually accompanied by excluding the atomic charges on the first MM group[22,23] and/or redistributing MM charges near the interface.[24–30] An alternative way to avoid an over-polarization of the QM wavefunction by point MM charges near the QM/MM interface is to reproduce them with Gaussian-blurred MM charges.[12,27,31,32]

QM/MM van der Waals interactions are usually treated empirically, where QM atoms are assigned vdW parameters based on atomic similarity, while MM atoms retain their vdW parameters from the MM force field. As such, QM/MM vdW interactions are not necessarily compatible with QM/MM electrostatics, especially during a chemical or physical process.[33,34] This has led to the development of several density-dependent QM/MM vdW models, which have yet to gain wide use.[35–42]

There are three general schemes for capturing the electrostatic interactions between QM and MM atoms. In the "continuous" scheme [Fig. 1(a)], the continuous QM electron density, $\rho(\boldsymbol{r})$, directly interacts with the MM charges, q_B,

$$E^{\text{elec, C}}_{\text{QM/MM}} = -\sum_{B\in\text{MM}}\int\frac{\rho(\boldsymbol{r})q_B}{|\boldsymbol{r}-\boldsymbol{r}_B|}d\boldsymbol{r} + \sum_{A\in\text{QM}}\sum_{B\in\text{MM}}\frac{Z_A q_B}{|\boldsymbol{r}_A-\boldsymbol{r}_B|}, \quad (1)$$

which was offset by the interactions between nuclear charges, Z_A, and MM charges. In a variation of this scheme,[43] one precomputes the MM electrostatic potentials on the grid positions in the QM region, $\phi(\boldsymbol{r})$, and at the QM nuclear positions, ϕ_A. Then, the energy becomes

$$E^{\text{elec, C}'}_{\text{QM/MM}} = -\int\rho(\boldsymbol{r})\phi(\boldsymbol{r})d\boldsymbol{r} + \sum_{A\in\text{QM}} Z_A\phi_A. \quad (2)$$

In the "surrogate" scheme [Fig. 1(b)], the QM electron density (and nuclei) is represented by surrogate charges, Q_A, and dipole moments, $\boldsymbol{\mu}_A$, and higher moments assigned to each QM atom. Together, these local multipole moments, $\mathscr{M}_{Am}$, interact with $\mathscr{F}_{Am}$, the local Taylor expansion of the MM electrostatic embedding potential,

$$\begin{aligned} E^{\text{elec, S}}_{\text{QM/MM}} &= \sum_{A\in\text{QM}}\sum_{m}\mathscr{M}_{Am}\mathscr{F}_{Am} \\ &= \sum_{A\in\text{QM}}\left[Q_A\phi_A - \boldsymbol{\mu}_A\cdot\mathscr{E}_A + \cdots\right], \end{aligned} \quad (3)$$

where the leading Taylor expansions include ϕ_A and $\mathscr{E}_A$, i.e., the local electrostatic potential and field (at the position of the Ath QM atom) due to all MM charges.

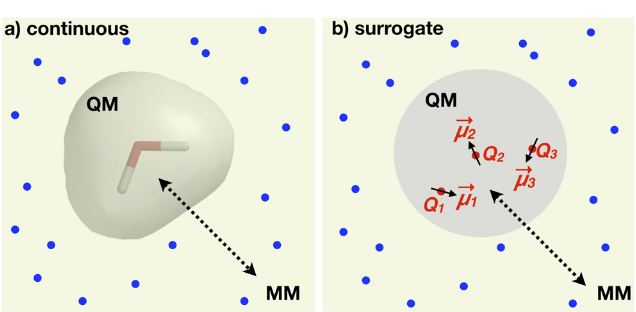

FIG. 1. Two schemes for describing the electrostatic interaction between QM and MM atoms: (a) a continuous QM electron density interacts with MM charges; (b) QM atoms are represented by surrogate multipoles in their interactions with MM charges.

Finally, the hybrid scheme combines a "continuous" description for the short-range (SR) QM/MM electrostatics and a "surrogate" description for the long-range (LR) QM/MM electrostatics,

$$E^{\text{elec, hybrid}}_{\text{QM/MM}} = E^{\text{elec, C}}_{\text{SR-QM/MM}} + E^{\text{elec, S}}_{\text{LR-QM/MM}}. \quad (4)$$

These three schemes have been implemented within various quantum chemistry and molecular mechanics programs as well as their interfaces.[23,30,57–68] The reader is referred to Ref. 69 for a complete review of these methodologies. Here, we shall only briefly discuss several models listed in Table I that are most relevant to this work.

The continuous scheme, where the continuous electron density interacts directly with the MM charges, is routinely used in *ab initio* QM/MM (ai-QM/MM) calculations on truncated systems. In the setup of these truncated systems, all MM atoms beyond a cutoff distance (typically around 15 Å–25 Å) from the QM region are removed, thus completely neglecting long-range QM/MM electrostatics or partially accounted for it through adding an implicit solvent environment. To maintain a proper boundary, a layer of MM atoms just within the cutoff are usually kept at fixed positions.

Within the continuous scheme, one essentially represents the MM embedding electrostatic potential in the basis of atomic orbitals $(\chi_\mu, \chi_\nu, \ldots)$,

$$V^{\text{MM}}_{\mu\nu} = \int\chi_\mu(\boldsymbol{r})\left[\sum_{B\in\text{MM}}\frac{q_B}{|\boldsymbol{r}-\boldsymbol{r}_B|}\right]\chi_\nu(\boldsymbol{r})\,d\boldsymbol{r}, \quad (5)$$

and uses it as part of the one-electron effective Hamiltonian (a.k.a. Fock matrix) to converge SCF energy. Unfortunately, such a continuous scheme becomes infeasible for extended systems because the cost of evaluating one-electron integrals in Eq. (5) (as well as their nuclear derivatives for obtaining ai-QM/MM energy gradient) increases linearly with the number of MM atoms.

In contrast, the alternative continuous scheme in Eq. (2) would allow us to retain a continuous description of QM electron density in the evaluation of long-range electrostatics of extended systems. For instance, in the ambient-potential composite Ewald method (CEw) from the work of Giese and York,[43] the long-range MM electrostatic potential was first computed on a rectangular grid in the QM region using the PME method. This potential was subsequently interpolated to the atom-centered Lebedev grid position for a numerical integration with the QM basis functions and electron density. A related Ewald-based approach was employed by Sanz–Navarro and coworkers in the SIESTA/Amber interface[58] and more recently by Kawashima, Ishimura, and Shiga.[70]

The surrogate schemes are commonly used in semi-empirical QM/MM (se-QM/MM) calculations. For instance, Cui and coworkers studied the interaction of Mulliken charges of QM atoms with MM charges in their SCC-DFTB/CHARMM calculations.[45,46] In ai-QM/MM calculations, Ferré and Ángyán proposed the use of ESP charges or charges/dipoles of QM atoms in the computation of ai-QM/MM electrostatic energy.[44] This method is current available within the MolCAS program[71–73] and through a Gaussian16/Tinker interface.[74–77]

TABLE I. A partial list of QM/MM electrostatic models.

Schemes	Methods	Systems	Description	References
Continuous	Ambient-potential composite Ewald (CEw)	se-QM/MM-PBC and ai-QM/MM-PBC	PME calculation of long-range MM electrostatic potentials on QM grid, $\phi(\boldsymbol{r})$, in Eq. (2)	Giese and York[43]
Surrogate	Electrostatic potential fitted operator (ESPF)	ai-QM/MM cluster	ESP charges/dipoles of QM atoms interacting with MM charges	Ferré and Ángyán[44]
	QM/MM-Ewald	se-QM/MM-PBC	Mulliken charges of QM atoms interacting with MM charges	Cui *et al.*[45,46]
Hybrid	QM/MM-PBC	ai-CP-QM/MM	Multipole moments of the QM region used in long-range electrostatics	Rothlisberger *et al.*[47]
	QM/MM-Ewald	se-QM/MM-PBC	Mulliken charges of QM atoms used in Ewald summation of long-range electrostatics	Nam, Gao, and York[48]
	QM/MM-PME	se-QM/MM-PBC	Mulliken charges of QM atoms used in PME calculation of long-range electrostatics	Walker, Crowley, and Case[49]
	QM/MM-Ewald	ai-QM/MM-PBC	ChElPG charges of QM atoms used in Ewald summation of long-range electrostatics	Herbert *et al.*[50,51]
	Dual focal ai-QM/MM-PME	ai-QM/MM-PBC	ESP charges of QM atoms used in PME calculation of long-range electrostatics	Zhou, Wang, Zhang[52]
	QM(LREC)/MM(PME)	ai-QM/MM-PBC	LREC function used to smooth transition between short-range and long-range electrostatics	Cisneros *et al.*[53,54]
	Gen-Ew	ai-QM/MM-PBC	PBC potential represented by virtual charges on a sphere	Vasilevskaya and Thiel[55]
	ESPC and ESPCD	ai-QM/MM cluster	Long-range MM potential represented by virtual charges on ESP grid	Pan, Rosta and Shao[56]
	QM/MM with augmentary charges (QM/MM-AC)	ai-QM/MM cluster ai-QM/MM-PBC	Long-range MM potential represented by augmentary charges on inner MM atoms	This work

The hybrid methods are widely adopted in se-QM/MM calculations on systems with a periodic boundary condition (PBC). The use of Mulliken charges of QM atoms to obtain the PBC correction of long-range se-QM/MM electrostatics was first proposed by Nam, Gao, and York.[48] While they obtained QM-MM and QM-QM PBC corrections using Ewald summation, Walker, Crowley, and Case calculated these corrections using the more efficient PME algorithm.[49]

Several hybrid methods have also been implemented for ai-QM/MM-PBC calculations. In an early implementation by the Rothlisberger group, the QM region was represented by its total charge, dipole, and quadrupole in the long-range QM/MM electrostatic interaction.[47] Subsequent implementations employed surrogate atomic charges for the QM region. Instead of using the Mulliken population scheme, which are known to be sensitive to the basis sets,[50] more stable schemes, such as ChElPG charges and other electrostatic potential (ESP) derived charges,[78–81] were adopted by Herbert, Zhang, and other groups.[50–52]

When using ESP-derived charges to represent the QM region in the long-range QM/MM electrostatics, one is essentially projecting the outer MM charges (as well as QM and MM charges in image cells in the case of a PBC system) onto the ESP grid for the charge fitting. An explicit charge projection (onto a sphere) of this kind was carried out in Vasilevskaya and Thiel's Gen-Ew model, which was built upon earlier Spherical Solvent Boundary Potential (SSBP),[82,83] Generalized Solvent Boundary Potential (GSBP),[84–86] and Solvated Macromolecular Boundary Potential (SMBP) methods.[87,88]

The accuracy of such charge projections was studied in an earlier work by some of the authors.[56] In the ESP-charge-based (ESPC) model, for example, the outer MM charges (beyond a cutoff distance of 10 Å) were interacted with ESP charges of QM atoms, which amounted to a projection of outer MM charges onto the Merz–Kollman grid (i.e., points on four layers of vdW surfaces).[80] In the ESP-charge-and-dipole-based (ESPCD) model, ESP dipoles on QM atoms were also included to further improve the accuracy (in terms of reproducing the electrostatic energy from the continuous scheme). When combined with the LREC function from the work of Cisneros and co-workers[53,54] for smoothing the transition between inner and outer MM regions, the ESPCD model could reproduce the total QM/MM energy within 0.1 kcal mol^{-1} and TDDFT/MM excitation energies within 0.001 eV, both compared to reference values of the same test systems within a large number of image cells.

Notwithstanding the high accuracy of the ESPCD model, it is not optimal to project outer MM charges onto the four layers of vdW surfaces, where the grid points are much closer to the QM region than the outer MM charges to be represented. More importantly, with each update of the geometry of the QM region, new grid points might appear, while some existing grid points might vanish, making it non-trivial to maintain a continuous PES surface and thus to compute the analytical energy gradient.

In this work, a simpler charge projection scheme is proposed, where outer MM charges are projected onto the position of inner MM atoms. Effectively, each inner MM atom receives an augmentary charge due to this projection. This QM/MM model with augmentary charges on inner MM atoms, which will be called "QM/MM-AC," is described in Sec. II. Its performance in ground-state QM/MM calculations is shown in Sec. III. Conclusions are drawn in Sec. IV.

II. METHODS

A. Hybrid QM/MM electrostatics

In a QM/MM calculation with electrostatic embedding, the total potential energy can be expressed as

$$\begin{aligned} E &= E_{\rm QM} + E_{\rm QM/MM} + E_{\rm MM} \\ &= E_{\rm QM} + E^{\rm elec}_{\rm QM/MM} + E^{\rm vdw}_{\rm QM/MM} + E^{\rm bound}_{\rm QM/MM} + E_{\rm MM}, \end{aligned} \tag{6}$$

where $E_{\rm QM}$ and $E_{\rm MM}$ are energies for the QM and MM subsystems, respectively. $E_{\rm QM/MM}$ is the coupling term between them, which can be further divided into electrostatic ($E^{\rm elec}_{\rm QM/MM}$), van der Waals ($E^{\rm vdw}_{\rm QM/MM}$), and the covalent bonding ($^{\rm bound}_{\rm QM/MM}$) terms. The $^{\rm bound}_{\rm QM/MM}$ is applied when one or more covalent bonds are cut at the QM–MM boundary.

As we mentioned in the Introduction, in typical QM/MM calculations, $E^{\rm vdw}_{\rm QM\text{-}MM}$ remains to be treated at the MM level using empirical parameters. If QM/MM vdW interactions are described empirically and a fixed-charge model is used for the MM subsystem, two energy terms ($E^{\rm vdw}_{\rm QM/MM}$ and $E_{\rm MM}$) in Eq. (6) can be decoupled from the QM calculations. The actual form of $E_{\rm QM}$ is determined by the specific QM method chosen for the calculation.

In the hybrid scheme, as shown in Eq. (4), the QM/MM electrostatic energy (i.e., MM charges interacting with both the QM electron density and nuclei) is broken into the short-, long-range, and periodic boundary correction terms, as

$$E^{\rm elec,\,hybrid}_{\rm QM/MM} = E^{\rm elec,\,C}_{\rm SR\text{-}QM/MM} + E^{\rm elec,\,S}_{\rm LR\text{-}QM/MM} + E^{\rm elec,\,S}_{\rm CR\text{-}QM/QM}, \tag{7}$$

where the last term is only included under the periodic boundary conditions, as discussed below. For each MM atom, a continuous minimum distance function[89]

$$d_B^{\min} = \frac{\alpha}{\ln\left[\sum_{A\in QM}\exp\left(\frac{\alpha}{d_{AB}}\right)\right]} \tag{8}$$

is used to calculate its distance from the QM region, where d_{AB} is the distance between the Bth MM atom and the Ath QM atom and α is an adjustable parameter. For systems with a periodic boundary condition, the minimum image convention is applied in the calculation of d_{AB} and $d_B^{\min}$ values.

A distance-based partitioning of MM atoms into inner and outer MM regions is rather straightforward. All MM atoms with a $d_B^{\min}$ value smaller than a cutoff distance $r_{\rm off}$ will fall into the inner MM region, and the interaction of their charges $q_B^<$ with the QM region will be called short-range QM/MM electrostatics and computed using Eq. (1). All other MM atoms in the center cell, as well as MM and QM charges in the image cells, will be assigned to the outer MM region, and their charges $q_B^>$ interact with the QM region through the "surrogate" model in Eq. (3), thus accounting for long-range QM/MM electrostatics.

However, such a sharp boundary between inner and outer MM regions causes a discontinuity in the total energy and gradient when a MM atom crosses the boundary. Therefore, we need to employ a switch function, $S(d_B^{\min})$, to smooth the transition between the two types of interactions at the cutoff distance. This function would decay smoothly from 1 to 0, when $d_B^{\min}$ increases from 0 to $r_{\rm off}$; Sec. II E shows four switch functions considered in the work. Thus, $q_B^<$ charge of each inner MM atom will be divided into two parts,

$$q_B^< = q_B^{<,C} + q_B^{<,S}, \tag{9}$$

$$q_B^{<,C} = S(d_B^{\min})q_B^<, \tag{10}$$

$$q_B^{<,S} = \left(1 - S(d_B^{\min})\right)q_B^<, \tag{11}$$

where only $q_B^{<,C}$ charges interact explicitly with a continuous QM electron density [see Fig. 2(a)]. Meanwhile, as shown in Fig. 2(b), $q_B^{<,S}$ charges are combined with outer MM charges (and QM charges in image cells) in the evaluation of long-range QM/MM electrostatics.

B. Charge projection for long-range QM/MM electrostatics

For the long-range QM/MM electrostatics in Eq. (3), the QM atoms are represented by $\mathscr{M}_{Am}$, a "surrogate" set of multipoles centered on the QM atom sites. These multipoles are fitted to reproduce ϕ_s, the electrostatic potential due to QM electron density and nuclei, on each point s of a pre-defined grid,

$$\sum_A\sum_m \mathscr{K}_{Am,s}\mathscr{M}_{Am} = \phi_s, \tag{12}$$

where $\mathscr{K}_{Am,s}$ is the interaction tensor between the QM atomic multipole moments $\mathscr{M}_{Am}$ and a unit charge on the grid point s. By employing singular-value decomposition (SVD) to invert the interaction tensor, we can obtain the ESP multipoles according to

$$\mathscr{M}_{Am} = \sum_s\left(\boldsymbol{\mathscr{K}}^{-1}\right)_{Am,s}\phi_s. \tag{13}$$

Based on Eq. (3), the long-range QM/MM electrostatic energy is

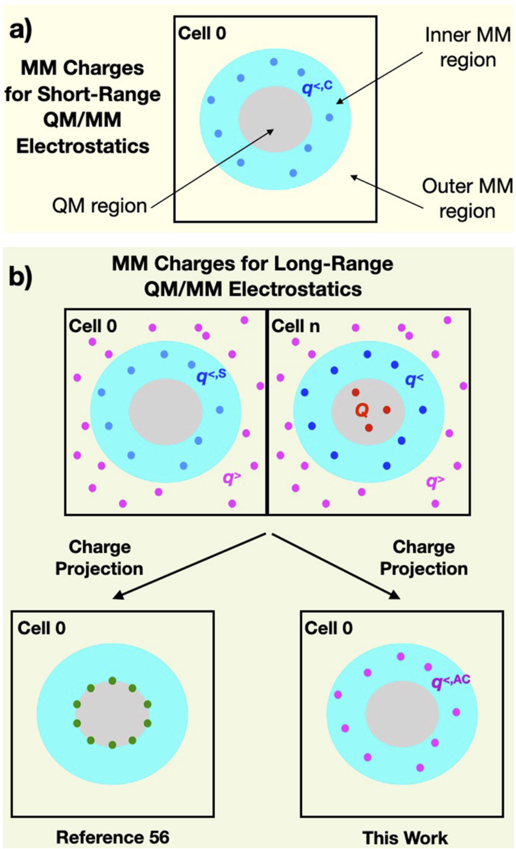

FIG. 2. Partitioning of the MM charges: (a) smoothed charges from inner MM atoms in Eq. (10) to be interacted with a continuous QM electron density in the short-range QM/MM electrostatics; (b) projected charges to be used to compute long-range QM/MM electrostatics. The projected charges represent (i) outer MM charges and remaining inner MM charges in Eq. (11) in the central cell and (ii) all QM and MM charges in image cells. Projected charges are located on either ESP fitting surfaces (Ref. 56) or inner MM atoms (this work). While our scheme [as shown in the last panel in 9b] resembles a truncated model, it should be noted that $q^{<,AC}$ charges capture the effect of long-range electrostatics.

$$E^{\text{elec, S}}_{\text{LR-QM/MM}} = \sum_{A\in\text{QM}}\sum_{m}\mathscr{M}_{Am}\mathscr{F}^{>}_{Am}, \tag{14}$$

where $\mathscr{F}^{>}_{Am}$ refers to the local Taylor expansion of the MM electrostatic embedding potential due to outer MM charges ($q^{>}_{B}$), QM charges in image cells ($Q^{n\neq 0}_{A}$), as well as $q^{<,S}_{B}$, which is the long-range portion of inner MM charges defined in Eq. (11). In actual implementation, one can calculate the $\mathscr{F}^{>}_{Am}$ values using Coulomb's law for non-PBC calculations and Ewald or PME for PBC calculations. Furthermore, $q^{<,C}_{B}$-QM charge interactions are subjected to the exclusion rules. Namely, for each QM atom site A, the contributions to $\mathscr{F}^{>}_{Am}$ from $q^{<,C}_{B}$, along with the contributions from the other QM charges (from the center cell for PBC calculations), and typically "MM1" atoms are excluded.[22,23] In our implementation, QM atoms in image cells adopt fixed charges, which can be pre-determined by following the standard protocol to fit MM partial charges for the corresponding force field. In this way, the generalized long-range MM embedding potential $\mathscr{F}^{>}_{Am}$ does not depend on the QM electron density, which is similar to the CEw method.[43]

By substituting $\mathscr{M}_{Am}$ in Eq. (13) into the energy expression in Eq. (14), we get

$$\begin{aligned} E^{\text{elec, S}}_{\text{LR-QM/MM}} &= \sum_{A}\sum_{m}\left[\sum_{s}\left(\boldsymbol{\mathscr{K}}^{-1}\right)_{Am,s}\phi_s\right]\mathscr{F}^{>}_{Am} \\ &= \sum_{s}\left[\sum_{A}\sum_{m}\mathscr{F}^{>}_{Am}\left(\boldsymbol{\mathscr{K}}^{-1}\right)_{Am,s}\right]\phi_s \\ &= \sum_{s} q_s\phi_s, \end{aligned} \tag{15}$$

where q_s are the projected charges on the ESP grid,

$$q_s = \sum_{A}\sum_{m}\mathscr{F}^{>}_{Am}\left(\boldsymbol{\mathscr{K}}^{-1}\right)_{Am,s}. \tag{16}$$

Equation (15) indicates that the long-range QM/MM electrostatic energy can also be viewed as the interaction energy between projected charges and QM electron density and nuclei.

This leads to an alternative way adopted in this work for handling the long-range QM/MM electrostatics, which was also employed in our previous ESPC and ESPCD models.[56] Instead of explicitly computing ESP-based multipoles for QM atoms, we will compute projected charges on the ESP grid points according to Eq. (16) and use these virtual charges to interact with and polarize QM electron density.

Traditionally, ESP charges are fitted by using grid points on the Merz–Kollman[80] or rectangular grid (e.g., CHELP and CHELPG)[78] outside the QM region. In our previous work, the Merz–Kollman grid (points on four layers of vdW surfaces) was employed. However, in our hybrid scheme, the long-range electrostatic embedding potential arises partially from $q^{<,S}_{B}$ from inner MM atoms (due to the use of the switching function). This contribution can be rather substantial because these virtual surface charges are located closer to the QM region than outer MM atoms. When simulating condensed-phase systems, to minimize the error (caused by our charge projection) in the interaction energy between those charges and the QM electron density, *a better way is to use the inner MM atom positions as the "grid" for charge fitting and projection.* Since only the interatomic distances between the QM and inner MM atoms were involved in the fitting, the resulting projection is naturally translationally and rotationally invariant. This is different from the use of grids on vdW surfaces, which does not maintain rotational invariance, or rectangular grids, which retains neither translational nor rotational invariance.

Henceforth, in this work, we will use this special "grid" to obtain projected charges that augment $q^{<,C}_{B}$ charges on inner MM atoms. We will therefore refer q_s in Eq. (16) as $q^{<,AC}_{B}$, where "AC" stands for "augmentary charges," and our overall QM/MM electrostatic scheme as "QM/MM-AC." To maintain a smooth potential energy surface at the cutoff distance, we can also use a weighting function w_B to scale the Coulomb interaction tensor, $(\boldsymbol{W}\mathscr{K})_{Am,B} = w_B\mathscr{K}_{Am,B}$, in the calculation of the projected charges,

$$q_B^{<,\mathrm{AC}} = \sum_A \sum_m \mathscr{F}_{Am}^{>} \left((\boldsymbol{W}\mathscr{K})^{-1} \right)_{Am,B} w_B, \tag{17}$$

to ensure that the projected charges vanish smoothly at the cutoff distance. In general, any weighting function that decays smoothly to zero at the cutoff distance can be used as w_B. In our work, for the sake of convenience, the same smoothening function in Sec. E is employed.

As an added benefit of our hybrid QM/MM-AC scheme, its support within a QM package can be trivial. Once the augmentary charges are computed by a MM package or QM/MM interface, one just needs to add the scaled and augmentary charges on inner MM atoms, $q_B^{<,\mathrm{C}} + q_B^{<,\mathrm{AC}}$, and include them as "external" point charges in QM calculations. In doing so, no extra modification is needed for the QM packages. The overall workflow is summarized in Algorithm 1.

C. QM–QM image electrostatic correction

For QM/MM-AC calculations under periodic boundary conditions, the generalized long-range MM embedding potential $\mathscr{F}_{Am}^{>}$ also includes contributions from the reference charges of QM atoms in all the image cells, which results in a double counting of the QM-QM image electrostatics. Instead of subtracting half of the QM-QM image electrostatics as done in Ref. 48,

$$E_{\text{CR-QM/QM}}^{\text{elec, S}} = -\frac{1}{2} \sum_{A \in \text{QM}} \sum_m \mathscr{M}_{Am} \mathscr{F}_{Am}^{\text{QM-image}}, \tag{18}$$

we followed the CEw method[43] and calculated the correction using

$$E_{\text{CR-QM/QM}}^{\text{elec, S}} = -\frac{1}{2} \sum_{A \in \text{QM}} q_A^{\text{ref}} \phi_A^{\text{QM-image}}, \tag{19}$$

where $\phi_A^{\text{QM-image}}$ is the electrostatic potential on the QM atom site A from the reference charges of the QM atoms from all the image cells, and can be calculated using the standard Ewald method efficiently. For non-PBC calculations, this correction term is not needed.

ALGORITHM 1. Workflow for computing QM/MM-AC electrostatic embedding potential.

1 Get $q_B^{<,\mathrm{C}}$, short-range portion of inner MM charges [Eq. (10)]
2 Get $q_B^{<,\mathrm{S}}$, long-range portion of inner MM charges [Eq. (11)]
3 **if** *PBC* **then**
4 Call helPME to compute, $\mathscr{F}_{Am}^{>}$, long-range electrostatic potential (and field, if necessary) due to $q_B^{<,\mathrm{S}}$, outer MM charges $q_B^{>}$, and all MM and (fixed-value) QM charges in the image cells
5 **else**
6 Compute, $\mathscr{F}_{Am}^{>}$, long-range electrostatic potential (and field, if necessary) due to $q_B^{<,\mathrm{S}}$ and outer MM charges $q_B^{>}$
7 **end**
8 Get $q_B^{<,\mathrm{AC}}$, augmentary charges on inner MM atoms from $\mathscr{F}_{Am}^{>}$ through charge projection [Eq. (17)]
9 Compute the total MM electrostatic embedding potential, $V_{\mu\nu}^{\mathrm{MM}}$, due to both $q_B^{<,\mathrm{C}}$ and $q_B^{<,\mathrm{AC}}$ [Eq. (5)]
10 Perform QM calculations using the MM embedding potential

D. Analytic gradient

As shown in Fig. 4, it is beneficial to use inner MM atom sites as the ESP grid for fitting QM atomic multipole moments and thus as sites for projecting outer MM charges. The accuracy of computed energies (at the same cutoff) is shown in Sec. III A, when compared to the use of points on vdW surfaces for multipole fitting and charge projection. As a result, our target accuracy (0.1 kcal mol^{-1} in total energy) can be achieved by only using ESP charges on the QM atoms, making it much easier to implement the analytical gradient. Hence, in the remainder of this subsection, only the ESP charges are used. Accordingly, the local Taylor expansion of the MM electrostatic embedding potential $\mathscr{F}_{Am}^{>}$ is truncated after the zeroth order,

$$\mathscr{F}_{Am}^{>} = \{\phi_A^{>}\}, \tag{20}$$

and the charge projection in Eq. (17) becomes

$$q_B^{<,\mathrm{AC}} = \sum_A \phi_A^{>} \left((\boldsymbol{W}\mathscr{K})^{-1} \right)_{A,B} w_B. \tag{21}$$

1. Analytic gradient on QM atoms

In the hybrid scheme, the gradient of the QM/MM electrostatic energy with respect to a Cartesian coordinate x_A of the QM atom A is

$$\frac{\partial E_{\text{QM/MM}}^{\text{elec, hybrid}}}{\partial x_A} = \frac{\partial E_{\text{SR-QM/MM}}^{\text{elec, C}}}{\partial x_A} + \frac{\partial E_{\text{LR-QM/MM}}^{\text{elec, S}}}{\partial x_A} + \frac{\partial E_{\text{CR-QM/QM}}^{\text{elec, S}}}{\partial x_A}, \tag{22}$$

where the last term is only required for QM/MM-AC calculations under periodic boundary conditions [see Eq. (19)].

The gradient of the short-range QM/MM electrostatics with respect to a Cartesian coordinate x_A of the Ath QM atom is

$$\begin{aligned}\frac{\partial E_{\text{SR-QM/MM}}^{\text{elec, C}}}{\partial x_A} = &\sum_{B \in \text{inner-MM}} \left[-\int \frac{\partial \rho(\boldsymbol{r})}{\partial x_A} \frac{1}{|\boldsymbol{r}-\boldsymbol{r}_B|} d\boldsymbol{r} - Z_A \frac{x_A - x_B}{R_{AB}^3} \right] q_B^{<,\mathrm{C}} \\ &+ \sum_{B \in \text{inner-MM}} \left[-\int \frac{\rho(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{r}_B|} d\boldsymbol{r} + \frac{Z_A}{R_{AB}} \right] \frac{\partial q_B^{<,\mathrm{C}}}{\partial x_A},\end{aligned} \tag{23}$$

where $R_{AB} = |\boldsymbol{r}_A - \boldsymbol{r}_B|$. The first term is the standard contribution from the external point charge to the QM gradient, which can be computed routinely by QM packages. Note that the density relaxation $\frac{\partial \rho(\boldsymbol{r})}{\partial x_A}$ does not have to contain the molecular orbital response contributions because the total QM/MM energy is variational to molecular orbital rotations. The second term in Eq. (23) arises from the external charge derivative $\partial q_B^{<,\mathrm{C}}/\partial x_A$ "interacting" with the electrostatic potential from the QM subsystem,

$$\phi_B^{\text{QM}} = -\int \frac{\rho(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{r}_B|} d\boldsymbol{r} + \frac{Z_A}{R_{AB}}, \tag{24}$$

at the MM site B. In regular QM/MM calculations, this term vanishes because $q_B^{<,\mathrm{C}}$ is typically fixed at the charge value defined in the force field. However, in the hybrid scheme, $\partial q_B^{<,\mathrm{C}}/\partial x_A$ is not necessarily zero because $q_B^{<,\mathrm{C}}$ is scaled by the weighting function that depends on the QM coordinates,

$$\frac{\partial q_B^{<,\mathrm{C}}}{\partial x_A} = \frac{\partial S(d_B^{\min})}{\partial x_A} q_B^{<}. \tag{25}$$

For long-range QM/MM electrostatics, the outer MM charges (and QM image charges) are described by the "surrogate" charges, $q_B^{<,\mathrm{AC}}$, as defined in Eq. (21). The corresponding QM gradient is similar to the "continuous" one, just with $q_B^{<,\mathrm{C}}$ replaced by $q_B^{<,\mathrm{AC}}$,

$$\frac{\partial E^{\mathrm{elec,\,S}}_{\mathrm{LR\text{-}QM/MM}}}{\partial x_A} = \sum_{B\in\mathrm{inner\text{-}MM}}\left[\int -\frac{\partial\rho(\boldsymbol{r})}{\partial x_A}\frac{1}{|\boldsymbol{r}-\boldsymbol{r}_B|}d\boldsymbol{r} - Z_A\frac{x_A-x_B}{R_{AB}^3}\right]q_B^{<,\mathrm{AC}} + \sum_{B\in\mathrm{inner\text{-}MM}}\phi_B^{\mathrm{QM}}\frac{\partial q_B^{<,\mathrm{AC}}}{\partial x_A}. \tag{26}$$

In the equation, the charge derivative is

$$\frac{\partial q_B^{<,\mathrm{AC}}}{\partial x_A} = \frac{\partial\phi_A^{>}}{\partial x_A}\left((\boldsymbol{W}\mathscr{K})^{-1}\right)_{A,B}w_B + \phi_A^{>}\frac{\partial\left((\boldsymbol{W}\mathscr{K})^{-1}\right)_{A,B}}{\partial x_A}w_B + \phi_A^{>}\left((\boldsymbol{W}\mathscr{K})^{-1}\right)_{A,B}\frac{\partial w_B}{\partial x_A}, \tag{27}$$

where $\partial\phi_A^{>}/\partial x_A$ in the first term is the gradient of the long-range MM embedding potential $\phi_A^{>}$ at QM site A, which is calculated using the PME method in this work. The second term,

$$\frac{\partial\left((\boldsymbol{W}\mathscr{K})^{-1}\right)_{A,B}}{\partial x_A} = \left(\frac{\partial(\boldsymbol{W}\mathscr{K})^{-1}}{\partial x_A}\right)_{A,B}, \tag{28}$$

can be calculated using the formula for the derivative of the pseudo-inverse of a matrix,[90]

$$\frac{\partial}{\partial x}\mathbf{A}^{-1} = -\mathbf{A}^{-1}\left(\frac{\partial}{\partial x}\mathbf{A}\right)\mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{A}^{-1\top}\left(\frac{\partial}{\partial x}\mathbf{A}^{\top}\right)(1-\mathbf{A}\mathbf{A}^{-1}) + (1-\mathbf{A}^{-1}\mathbf{A})\left(\frac{\partial}{\partial x}\mathbf{A}^{\top}\right)\mathbf{A}^{-1\top}\mathbf{A}^{-1}. \tag{29}$$

2. Analytic gradient on inner MM atoms

For gradient on inner MM atoms, the contribution from short-range QM/MM electrostatics is

$$\frac{\partial E^{\mathrm{elec,\,C}}_{\mathrm{SR\text{-}QM/MM}}}{\partial x_B} = \left[-\int \rho(\boldsymbol{r})\frac{x-x_B}{|\boldsymbol{r}-\boldsymbol{r}_B|^3}d\boldsymbol{r} + \sum_{A\in\mathrm{QM}}Z_A\frac{x_A-x_B}{R_{AB}^3}\right]q_B^{<,\mathrm{C}} + \phi_B^{\mathrm{QM}}\frac{\partial q_B^{<,\mathrm{C}}}{\partial x_B}, \tag{30}$$

where the first term is typically evaluated by computing the electrostatic field of QM electrons/nuclei at the position of the Bth MM atom and then scaling it by the charge value, $q_B^{<,\mathrm{C}}$, and the charge derivatives, $\frac{\partial q_B^{<,\mathrm{C}}}{\partial x_B}$ is computed in a similar way as Eq. (25).

The corresponding long-range QM/MM electrostatic contribution is

$$\frac{\partial E^{\mathrm{elec,\,S}}_{\mathrm{LR\text{-}QM/MM}}}{\partial x_B} = \left[-\int \rho(\boldsymbol{r})\frac{x-x_B}{|\boldsymbol{r}-\boldsymbol{r}_B|^3}d\boldsymbol{r} + \sum_{A\in\mathrm{QM}}Z_A\frac{x_A-x_B}{R_{AB}^3}\right]q_B^{<,\mathrm{AC}} + \phi_B^{\mathrm{QM}}\frac{\partial q_B^{<,\mathrm{AC}}}{\partial x_B}, \tag{31}$$

where the charge derivatives are

$$\frac{\partial q_B^{<,\mathrm{AC}}}{\partial x_B} = \sum_{A\in\mathrm{QM}}\frac{\partial\phi_A^{>}}{\partial x_B}\left((\boldsymbol{W}\mathscr{K})^{-1}\right)_{A,B}w_B + \sum_{A\in\mathrm{QM}}\phi_A^{>}\frac{\partial\left((\boldsymbol{W}\mathscr{K})^{-1}\right)_{A,B}}{\partial x_B}w_B + \sum_{A\in\mathrm{QM}}\phi_A^{>}\left((\boldsymbol{W}\mathscr{K})^{-1}\right)_{A,B}\frac{\partial w_B}{\partial x_B}, \tag{32}$$

and the first term involves

$$\frac{\partial\phi_A^{>}}{\partial x_B} = \frac{\partial}{\partial x_B}\left(\frac{q_B^{<,S}}{R_{AB}}\right) = q_B^{<,S}\frac{x_A-x_B}{R_{AB}^3} + \frac{\partial S(d_B^{\min})}{\partial x_B}\frac{q_B^{<}}{R_{AB}}. \tag{33}$$

3. Analytic gradient on outer MM atoms

Outer MM charges interact with the QM atoms only through long-range electrostatics,

$$\begin{aligned}\frac{\partial E^{\mathrm{elec,\,S}}_{\mathrm{LR\text{-}QM/MM}}}{\partial x_{B'}} &= \sum_{B\in\mathrm{inner\text{-}MM}}\phi_B^{\mathrm{QM}}\frac{\partial q_B^{<,\mathrm{AC}}}{\partial x_{B'}} \\ &= \sum_{A\in\mathrm{QM}}\frac{\partial\phi_A^{>}}{\partial x_{B'}}\sum_{B\in\mathrm{inner\text{-}MM}}\left((\boldsymbol{W}\mathscr{K})^{-1}\right)_{A,B}w_B\phi_B^{\mathrm{QM}} \\ &= \sum_{A\in\mathrm{QM}}\frac{\partial\phi_A^{>}}{\partial x_{B'}}q_A^{\mathrm{ESP}},\end{aligned} \tag{34}$$

where

$$q_A^{\mathrm{ESP}} = \sum_{B\in\mathrm{inner\text{-}MM}}\left((\boldsymbol{W}\mathscr{K})^{-1}\right)_{A,B}w_B\phi_B^{\mathrm{QM}} \tag{35}$$

is the QM ESP charges fitted using the inner MM atom sites as the weighted grid. Thus, this term arises from the electric field at outer MM atom positions from the QM ESP charges, which can be readily calculated using the PME method.

E. Smoothing functions

As in our previous work,[56] four smoothing functions were considered in this study, including the step function,

$$S^{\mathrm{Step}}(r) = \begin{cases}1, & r\leqslant r_{\mathrm{off}}\\ 0, & r > r_{\mathrm{off}},\end{cases} \tag{36}$$

the shift function,[91]

$$S^{\mathrm{Shift}}(r) = \begin{cases}(1-(r/r_{\mathrm{off}})^2)^2 & r\leqslant r_{\mathrm{off}},\\ 0, & r > r_{\mathrm{off}},\end{cases} \tag{37}$$

the switch function,[91]

$$S^{\mathrm{Switch}}(r) = \begin{cases}1, & r\leqslant r_{\mathrm{on}}\\ \dfrac{(r_{\mathrm{off}}^2-r^2)^2(r_{\mathrm{off}}^2+2r^2-3r_{\mathrm{on}}^2)}{(r_{\mathrm{off}}^2-r_{\mathrm{on}}^2)^3}, & r_{\mathrm{on}} < r\leqslant r_{\mathrm{off}}\\ 0, & r > r_{\mathrm{off}},\end{cases} \tag{38}$$

with r_{on} set to be $0.75r_{\mathrm{off}}$, and the long-range electrostatic correction (LREC) function,[53,54]

$$S^{\mathrm{LREC}}(r) = \begin{cases} 1 - \left[2\left(1 - \frac{r}{r_{\mathrm{off}}}\right)^3 - 3\left(1 - \frac{r}{r_{\mathrm{off}}}\right)^2 + 1\right]^2, & r \leqslant r_{\mathrm{off}} \\ 0, & r > r_{\mathrm{off}}. \end{cases} \tag{39}$$

These functions are employed in Eqs. (10) and (11) to split inner MM charges into continuous and surrogate portions. They are also used in Eq. (17) to define the weighting function w_B in the charge fitting.

III. IMPLEMENTATION AND COMPUTATION DETAILS

Our QM/MM-AC scheme was implemented and tested through a generic QM/MM interface named QMHub, which in our work connects the AMBER molecular mechanics software package[92] and a development version of the Q-CHEM 5.0 software package.[63] Specifically, QMHub collects atomic coordinates and charges from AMBER (which performs the MD sampling), divides the system into QM, inner, and outer MM regions, carries out the charge projection, and prepares Q-CHEM QM/MM input files. Subsequently, Q-CHEM performs a standard ai-QM/MM calculation with inner MM charges (scaled charges augmented by projected charges from outer atoms). Information is then sent back to QMHub, which computes forces on all MM atoms and sends them to AMBER for driving dynamics. The long-range electrostatic potential and electric field are calculated by PME using the helPME package developed by Simmonett (https://github.com/andysim/helpme). Below, we shall describe the three test systems used to validate our QM/MM-AC scheme.

A. Anionic oxyluciferin in luciferase

Neutral and anionic oxyluciferin (OLU and OLU$^-$) in both aqueous and enzyme environments were used in our previous work,[56] where several electrostatic models were compared. It was concluded that the ESPCD model, where ESP charges and dipoles were used to reproduce the electrostatic potentials on the Merz–Kollman grid, was needed to achieve the accuracy we targeted (0.1 kcal mol^{-1} in total energy).

In this work, we revisited the most challenging case from our previous study,[56] anionic oxyluciferin [OLU$^-$, Fig. 3(a)] in luciferase. Similar calculations were performed using the ESPC and ESPCD models but with inner MM atom sites as the ESP fitting grid. We considered two options for the ESP charge fitting: (a) each MM atom position was assigned an equal weight and (b) the weights for inner MM atom points were dependent on their distances to the nearest QM atoms. In the latter case, the same switching function, $S(d_B^{\min})$, which we used to partition the inner MM charges, was also adopted as the weighting function in Eq. (21) for ESP charge fitting.

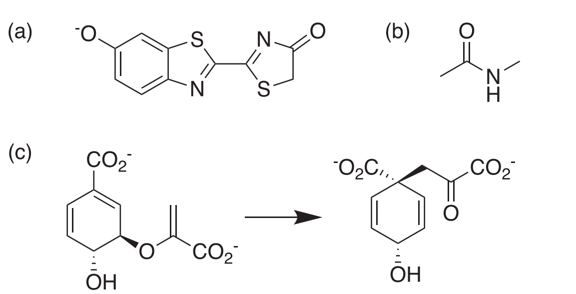

FIG. 3. Structures for (a) anionic oxyluciferin and (b) *N*-methylacetamide (NMA); (c) chorismate mutase catalyzed reaction.

All calculations were carried out using the same 100 configurations from a classical MD trajectory as in our previous work.[56] For the reference calculations, a very large supercell (with 123 image cells, ~20 000 000 atoms) was built from the original simulation box of 117 × 117 × 117 Å^3 to mimic a periodic system. The QM subsystem consisted of the OLU$^-$ molecule in the center cell and was described by density functional theory with the B3LYP functional[93–95] and 6-31+G* basis set.[96–98] The rest of the center cell and all the image cells, including the image OLU$^-$ molecules, became the MM subsystem and was represented by their partial atomic charges from the C36/CGenFF/TIP3P force fields.[99–102]

In QM/MM calculations using the ESPC or ESPCD models, the setup of the QM and MM subsystems, QM method, and MM charges were identical to the reference calculations. An atom-centered cutoff around the QM subsystem was used to divide the MM subsystem into inner and outer MM regions. Atom-based cutoffs were applied in cases where a switching function (shift, switch, or LREC) was used, while group-based cutoffs were employed when the step function was used to avoid artificial net charges in the near-field region.

B. Solvated NMA

Solvated *N*-methylacetamide [NMA, Fig. 3(b)], which is a model system for peptide, was used to check the energy conservation of our QM/MM-AC scheme in a microcanonical (NVE) simulation. Specifically, it contains one NMA molecule solvated in a cubic box containing 1661 TIP3P water molecules (~37 × 37 × 37 Å^3 after equilibration). QM/MM-AC NVE simulations were conducted under periodic boundary conditions using a modified version of SANDER.

The QM subsystem included only the NMA molecule, which was described at the B3LYP/6-31G* level of theory, whereas the solvent molecules constituted the MM subsystem. The van der Waals interactions between the QM and MM subsystems were modeled at the MM level using Lennard-Jones (LJ) potential, and the LJ parameters for NMA were taken from the ff14SB force field.[103] The RESP charges of NMA, which were obtained using the standard AMBER protocol, were used to represent the QM atoms in the image cells for the intercell QM–QM interactions. For ai-QM/MM electrostatics, we employed the QM/MM-AC algorithm, which adopts the LREC switch function in the ESPC electrostatic modeling with weighted inner MM positions as the ESP grid.

For comparison, we also performed MM MD simulations as well as se-QM/MM MD simulations using the built-in QM/MM functionality of SANDER with PM3 as the se-QM method. In all three NVE simulations, a time step of 0.5 fs was used for all the simulations. The PME method was employed to treat the electrostatic interactions, while the van der Waals interactions were truncated at a cutoff of 10 Å. The SHAKE algorithm was used to constrain all the bonds involving hydrogen atoms in the MM subsystem. The system was heated and equilibrated at the MM level at 300 K and 1 atm, and a 100 ps NVE production run was performed for each of the MM, se-QM/MM, and ai-QM/MM simulations.

C. Chorismate mutase

We calculated the free energy profile of the chorismate mutase reaction [Fig. 3(c)] using the QM/MM-AC scheme under periodic boundary conditions. The QM subsystem consisted of the substrate chorismate and was described by the B3LYP/6-31G* level of theory. Meanwhile, the MM subsystem included the enzyme, water solvent molecules, and Na^+ counter ions in the center cell, as well as all the atoms in the image cells (including the QM images), and was described by ff14SB/GAFF/TIP3P forces fields.[102–104] Langevin dynamics with a friction coefficient of 5 ps^{-1} was performed at 300 K, and a time step of 1 fs was used for the MD integration. The simulations were performed in the NVT ensemble with a box size of $76 \times 76 \times 76$ Å^3. The rest of setup was similar to the NMA simulation stated above.

The umbrella sampling technique[105] was used to estimate the free energy profile along the reaction coordinate, which was defined as the difference between the bond lengths of the breaking and forming bonds in this study. 40 windows were evenly distributed along the reaction coordinate ranged from −1.95 Å to 1.95 Å, and the force constant of the harmonic biasing potential was set to be 300 kcal mol^{-1} Å^{-2} for all the windows. For each window, 30 ps ai-QM/MM MD simulation was performed, and the snapshots were collected every 20 steps during the last 20 ps simulation, which resulted in 1000 snapshots in each window for further analysis. The Multistate Bennett acceptance ratio (MBAR)[106] as implemented in the pymbar package (https://github.com/choderalab/pymbar) was used to compute the free energy profile.

IV. RESULTS AND DISCUSSIONS

A. Accuracy of using MM atom sites as ESP fitting grid

For anionic oxyluciferin (OLU^-) in its enzyme environment, Fig. 4 showed the root-mean-square deviations (RMSD) in the QM/MM electrostatic and polarization energies from the theoretical reference values. Here, QM/MM permanent electrostatic energy referred to the interaction of oxyluciferin anion at its gas-phase electronic density with the extended MM electrostatic environment. QM/MM polarization energy, on the other hand, corresponded to the energy lowering due to the polarization of QM electron density by the MM charges. The deviations, as averaged over 100 different configurations for the system, were shown at different cutoff distances, r_{off}, ranging from 5 Å to 30 Å. Our objective was to identify embedding schemes that could produce accurate values (i.e., within 0.1 kcal mol^{-1} from the reference values) for both energies at a standard cutoff distance (~10 Å).

In general, the errors in QM/MM permanent and polarization energy are expected to decay rapidly with larger cutoff distances because more MM atoms were assigned to the inner MM region and interacted explicitly with the QM density. However, with the "Step" option, where no smoothing occurred at all at the cutoff distance, the energy errors oscillated with the cutoff distance in panels (a1), (a2), and (a4) of Fig. 4. In contrast, a steady decay in the energy errors was observed with three smoothing functions (LREC, switch, and shift). This reaffirmed the importance of using smoothing functions to ensure a smooth transition between inner and outer MM regions.

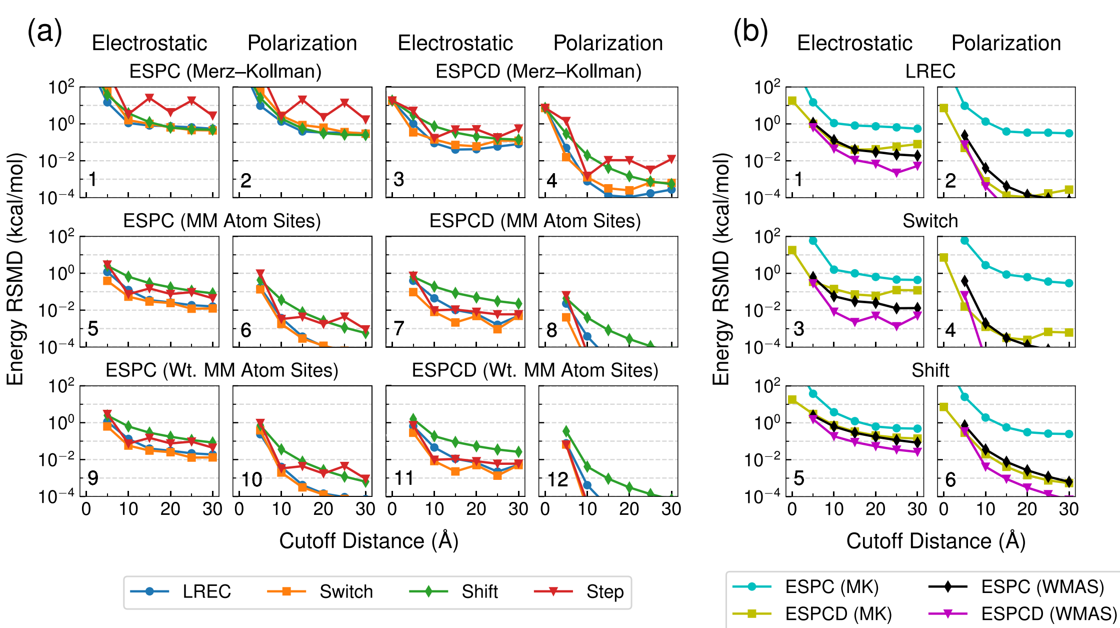

FIG. 4. Errors in ai-QM/MM permanent electrostatic and polarization energies (in kcal mol^{-1}) of anionic oxyluciferin in luciferase with different cutoff distances (in Å). (a) ESP charges (ESPC) and ESP charges/dipoles (ESPCD) were fitted to reproduce the long-range electrostatic potential (due to QM nuclei and electron density) on Merz–Kollman (MK) grid points on four layers of vdW surfaces [panels (a1)–(a4)], MM atomic sites [panels (a5)–(a8)], or weighted MM atomic sites [WMAS, panels (a9)–(a12), see Eq. (17)]. The ESPC-WMAS model in panels (a9) and (a10) with a 10 Å cutoff and the "LREC" switch function is the "QM/MM-AC" model recommended for use. (b) The energy errors from panels (a1) to (a12) are grouped with the LREC (b1) and (b2), switch (b3) and (b4), and shift (b5) and (b6) smoothing functions.

Among three smoothing functions, the "Shift" function displayed similar performance as "LREC" and "Switch" for the ESPC-MK model [panels (a1) and (a2)]. However, for all other models [panels (a3)–(a12)], "LREC" and "Switch" functions consistently led to lower errors than the "Shift" function. Thus, as discovered in our previous work,[56] "LREC" and "Switch" should be the preferred choices for the smoothing function in our QM/MM electrostatic calculations.

Among the six methods, ESPC-MK model [panels (a1) and (a2)] never achieved our target accuracy of 0.1 kcal mol^{-1}, even with a 30 Å distance cutoff. Hence, the original ESPC-MK model was not recommended.[56] In contrast, at a 10 Å distance cutoff, all other models with a "LREC" or "Switch" function already produced an error at or below 0.1 kcal mol^{-1} in QM/MM electrostatic energy [see panels (b1) and (b3)] and an error below 0.01 kcal mol^{-1} in QM/MM polarization energy [see panels (b2) and (b4)]. While these five models (ESPCD-MK, ESPC-MM, ESPCD-MM, ESPC-WMM, and ESPCD-WMM) all met our target accuracy, the ESPC-MM and ESPC-WMM models looked the most attractive because it is easier to formulate the analytical gradient of an ESPC model than that of an ESPCD model.

In panels (a9) and (a10), when the inner MM atom positions are weighted for charge fitting/projection in the ESPC-WMM model, the resultant results (at the logarithm scale) showed no substantial difference from those of the ESPC-MM model [panels (a5) and (a6)]. In the end, our favorite electrostatic embedding model is the ESPC-WMM model at a 10 Å cutoff (on the blue curves) in panels (a9) and (a10), which combines a "surrogate" ESP charge description for the QM region with weighted inner MM atom positions as the ESP fitting grid, 10 Å distance cutoff, and "LREC" smoothing function. We will refer to this model as the "QM/MM-AC" model in the remainder of this paper.

B. Accuracy of analytic energy gradient

Table II lists the QM/MM gradient values for a reactant configuration of the chorismate mutase reaction. Four QM atoms involved in the chemical reaction (C1, C5, O7, and C9), two inner MM atom (Arg90–N_ϵ; water-1057 oxygen), and one outer MM atom (water-440 oxygen) were considered as representative atoms in each region. The maximum difference between the analytical and numerical gradients was found to be 0.049 kcal mol^{-1} $Å^{-1}$. This confirmed that our analytical QM/MM gradient was properly implemented.

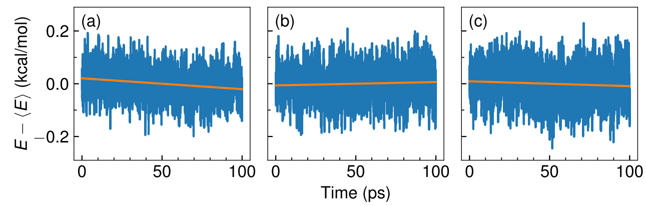

FIG. 5. Fluctuation in (a) total MM energy, (b) total se-QM/MM energy, and (c) total ai-QM/MM energy in NVE simulations of a NMA molecule solvated in a box of 1661 TIP3P water molecules. PM3 and B3LYP/6-31G* levels of theory were used in se-QM/MM and ai-QM/MM calculations, respectively.

C. Energy conservation in microcanonical MD simulations

As a further validation of the analytical gradient, the energy of a solvated NMA in NVE simulations was computed and shown in Fig. 5. In a pure MM simulation of the periodic system [Fig. 5(a)], the energy was found to drift by an average of −0.0004 kcal mol^{-1} per ps, while a drift rate of 0.0001 kcal mol^{-1} ps^{-1} was found from a se-QM/MM simulation with the PM3 model applied on the NMA molecule [Fig. 5(b)]. A comparable drift rate of −0.0002 kcal mol^{-1} ps^{-1} can be observed in Fig. 5(c) for our new QM/MM-AC model with a B3LYP/6-31G* level description for the NMA molecule.

D. Free energy profile for chorismate mutase

The free energy profile of the chorismate mutase reaction using ai-QM/MM under PBC was shown in Fig. 6. The free energy pro-

TABLE II. Comparison of analytical and numerical gradients for chorismate mutase (kcal mol^{-1} $Å^{-1}$). Three-point stencil and displacements of ±0.001 Å were used in the finite-difference calculation.

	Analytical			Numerical			Difference		
Atom	x	y	z	x	y	z	x	y	z
C1–CHO[a]	−11.903	−30.438	−25.330	−11.921	−30.422	−25.316	−0.018	0.016	0.014
C5–CHO	30.377	47.644	20.784	30.360	47.712	20.811	−0.017	0.068	0.027
O7–CHO[b]	−15.323	4.165	0.148	−15.314	4.162	0.148	0.009	−0.003	0.000
C9–CHO	7.003	31.950	1.070	6.999	31.992	1.072	−0.004	0.042	0.002
NE-ARG90[c]	37.012	−43.649	37.000	36.991	−43.665	37.049	−0.021	−0.016	0.049
O-WAT1057[d]	14.648	−0.955	−13.235	14.640	−0.954	−13.227	−0.008	0.001	0.008
O-WAT440[e]	36.766	8.161	4.189	36.746	8.160	4.187	−0.020	−0.001	−0.002

[a]C1–CHO and C5–CHO are the atoms of the forming bond during the chorismate mutase reaction.
[b]O7–CHO and C9–CHO are the atoms of the breaking bond during the chorismate mutase reaction.
[c]NE-ARG90 is the ϵ nitrogen of ARG90.
[d]O-WAT1057 is an arbitrary water oxygen between 9 Å and 10 Å from the QM region.
[e]O-WAT440 is an arbitrary water oxygen between 24 Å and 25 Å from the QM region.

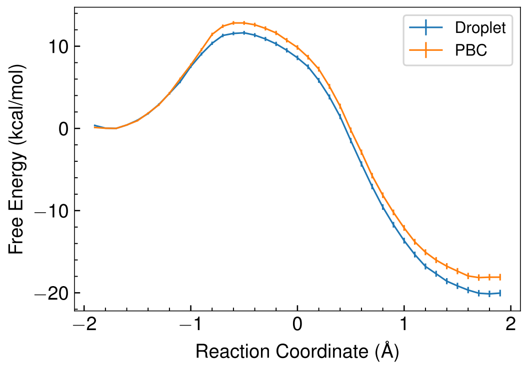

FIG. 6. Potential of mean force for the chorismate mutase reaction. The error bars show the uncertainties estimated by the MBAR method.

file of the same reaction using a droplet model (with a cutoff distance of 25 Å) from our previous study[107] was included for a comparison. While the two profiles are qualitatively very similar, the differences in the free energy barrier and reaction free energy are non-negligible (>1 kcal mol^{-1}). The difference mainly arose from two sources. The first source is that in the droplet model, the atoms beyond 25 Å were deleted from the system, and thus, the long-range QM-MM electrostatic interactions were missing. In contrast, our new QM/MM-AC model captured the electrostatic potential from all MM atoms. The second source is the periodic boundary condition, which avoids arbitrary constraints/restraints near the outer shell of the droplet, thus removing a bias toward specific boundary configurations and ensuring a proper sampling of all protein/solvent conformations.

E. Timing

For the chorismate mutase reaction, the aggregated wall time for 30 ps per window for 40 windows of umbrella sampling was 3822.4 h on one 20-core node (2.3 GHz Intel Xeon E5-2650 v3 processors), which arose mainly from the QM/MM energy and force calculations using the Q-CHEM software package. In contrast, the droplet model took 5687.6 h of simulation time. The 30% reduction in computer time was made possible by having fewer MM charges interact explicitly with the QM electron density.

V. CONCLUSIONS

In QM/MM calculations of condensed-phase reactions (as well as photochemical processes), it is still common to use the cutoff model, where all MM atoms beyond a distance cutoff (~15 Å–25 Å beyond the macromolecule) are removed. Then, to prevent the solvent molecules from "flying away," one can either apply a restraining potential to keep all atoms within the outer boundary or restrain an outer shell of MM atoms (usually 5 Å–10 Å thin) around their initial positions. Such a cutoff model artificially restricts the motion of MM atoms near the outer boundary, while ignoring long-range electrostatics (or approximating it with a continuum medium model). The effect of these restrictions/approximations, which can sometimes be rather substantial (as shown in Fig. 6), is hard to predict in a practical QM/MM calculation. As a result, in order to get reliable free energy results, one usually performs parallel simulations from different initial conformations,[108] thus further increasing the computational cost.

This problem can, in principle, be resolved by carrying out QM/MM calculations using a much larger cutoff distance or, more appropriately, a periodic boundary condition. This has inspired the development of many QM/MM-PBC models mentioned earlier in the Introduction. *Our QMMM-AC model builds upon previous QM/MM-PBC models and offers an alternative but effective protocol for handling QM/MM electrostatics of both PBC and non-PBC systems.* It allows us to separate inner MM and outer MM regions with a cutoff distance of 10 Å and projects outer MM charges onto inner MM atom positions. The model is accurate, leading to an error less than 0.1 kcal mol^{-1} in the total QM/MM energy. It is cost-effective, with a computational time demand even lower than typical cutoff models (with a ~25 Å cutoff). It is also compatible with most QM packages, requiring only an augmentation to the charge values of inner MM atoms passed to the QM program.

An efficient QM/MM electrostatic evaluation method will facilitate QM/MM calculations of systems with either a periodic boundary condition or a cluster with a large number of MM atoms. In addition, it will expedite QM/MM calculations with a single QM region embedded within an ensemble of MM environments, such as average solvent/macromolecule potential methods,[109–112] free energy gradient method,[113–115] and QM/MM-MFEP method.[6]

Finally, we note that this work focuses entirely on the treatment of long-range QM/MM electrostatics. An accurate and efficient long-range QM/MM electrostatic model, like the QM/MM-AC model proposed in this work, should be combined with an appropriate treatment of short-range electrostatics. As mentioned in the Introduction, for cases where the QM region is covalently linked to the MM region, MM charges near the covalent interface should be redistributed and/or Gaussian-blurred to avoid an over-polarization of the QM wavefunction.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health through Grant No. R43GM133270 to E.E., K.N., and Y.S. Y.S. was also supported by the National Institutes of Health (Grant No. R01GM135392), the Oklahoma Center for the Advancement of Science and Technology (Grant No. HR18-130), and the Office of the Vice President of Research and the College of Art and Sciences at the University of Oklahoma (OU). K.N. was also supported by the National Institutes of Health (Grant No. R01GM132481). A.C.S. was supported by the intramural research program of the National Heart, Lung, and Blood Institute. E.R. acknowledges funding from EPSRC (Grant No. EP/R013012/1) and ERC (Project No. 757850 BioNet).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] U. C. Singh and P. A. Kollman, "A combined *ab initio* quantum mechanical and molecular mechanical method for carrying out simulations on complex molecular systems: Applications to the $CH_3Cl + Cl^-$ exchange reaction and gas phase protonation of polyethers," J. Comput. Chem. **7**, 718–730 (1986).

[2] J. Gao and X. Xia, "A priori evaluation of aqueous polarization effects through Monte Carlo QM-MM simulations," Science **258**, 631–635 (1992).

[3] J. Gao and D. G. Truhlar, "Quantum mechanical methods for enzyme kinetics," Annu. Rev. Phys. Chem. **53**, 467–505 (2002).

[4] D. Riccardi, P. Schaefer, Y. Yang, H. Yu, N. Ghosh, X. Prat-Resina, P. König, G. Li, D. Xu, H. Guo, M. Elstner, and Q. Cui, "Development of effective quantum mechanical/molecular mechanical (QM/MM) methods for complex biological processes," J. Phys. Chem. B **110**, 6458–6469 (2006).

[5] H. Lin and D. G. Truhlar, "QM/MM: What have we learned, where are we, and where do we go from here?," Theor. Chem. Acc. **117**, 185–199 (2007).

[6] H. Hu and W. Yang, "Free energies of chemical reactions in solution and in enzymes with *ab initio* quantum mechanics/molecular mechanics methods," Annu. Rev. Phys. Chem. **59**, 573–601 (2008).

[7] H. M. Senn and W. Thiel, "QM/MM methods for biomolecular systems," Angew. Chem., Int. Ed. **48**, 1198–1229 (2009).

[8] M. W. van der Kamp and A. J. Mulholland, "Combined quantum mechanics/molecular mechanics (QM/MM) methods in computational enzymology," Biochemistry **52**, 2708–2728 (2013).

[9] M. Karplus, "Development of multiscale models for complex chemical systems: From H + H_2 to biomolecules (nobel lecture)," Angew. Chem., Int. Ed. **53**, 9992–10005 (2014).

[10] A. Warshel, "Multiscale modeling of biological functions: From enzymes to molecular machines (Nobel lecture)," Angew. Chem., Int. Ed. **53**, 10020–10031 (2014).

[11] S. Ahmadi, L. Barrios Herrera, M. Chehelamirani, J. Hostaš, S. Jalife, and D. R. Salahub, "Multiscale modeling of enzymes: QM-cluster, QM/MM, and QM/MM/MD: A tutorial review," Int. J. Quantum Chem. **118**, e25558 (2018).

[12] D. Das, K. P. Eurenius, E. M. Billings, P. Sherwood, D. C. Chatfield, M. Hodošček, and B. R. Brooks, "Optimization of quantum mechanical molecular mechanical partitioning schemes: Gaussian delocalization of molecular mechanical charges and the double link atom method," J. Chem. Phys. **117**, 10534–10547 (2002).

[13] V. Théry, D. Rinaldi, J.-L. Rivail, B. Maigret, and G. G. Ferenczy, "Quantum mechanical computations on very large molecular systems: The local self-consistent field method: Quantum mechanical computations," J. Comput. Chem. **15**, 269–282 (1994).

[14] N. Ferré, X. Assfeld, and J.-L. Rivail, "Specific force field parameters determination for the hybrid *ab initio* QM/MM LSCF method: The hybrid *ab initio* QM/MM LSCF method," J. Comput. Chem. **23**, 610–624 (2002).

[15] J. Gao, P. Amara, C. Alhambra, and M. J. Field, "A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations," J. Phys. Chem. A **102**, 4714–4721 (1998).

[16] J. Pu, J. Gao, and D. G. Truhlar, "Generalized hybrid-orbital method for combining density functional theory with molecular mechanicals," Chem. Phys. Chem. **6**, 1853–1865 (2005).

[17] Y. Zhang, T.-S. Lee, and W. Yang, "A pseudobond approach to combining quantum mechanical and molecular mechanical methods," J. Chem. Phys. **110**, 46–54 (1999).

[18] Y. Zhang, "Improved pseudobonds for combined *ab initio* quantum mechanical/molecular mechanical methods," J. Chem. Phys. **122**, 024114 (2005).

[19] D. M. Philipp and R. A. Friesner, "Mixed *ab initio* QM/MM modeling using frozen orbitals and tests with alanine dipeptide and tetrapeptide," J. Comput. Chem. **20**, 1468–1494 (1999).

[20] V. Kairys and J. H. Jensen, "QM/MM boundaries across covalent bonds: A frozen localized molecular orbital-based approach for the effective fragment potential method," J. Phys. Chem. A **104**, 6656–6665 (2000).

[21] Y. Shao and J. Kong, "YinYang atom: A simple combined *ab initio* quantum mechanical molecular mechanical model," J. Phys. Chem. A **111**, 3661–3671 (2007).

[22] H. L. Woodcock, M. Hodošček, A. T. B. Gilbert, P. M. W. Gill, H. F. Schaefer, and B. R. Brooks, "Interfacing Q-CHEM and CHARMM to perform QM/MM reaction path calculations," J. Comput. Chem. **28**, 1485–1502 (2007).

[23] A. W. Götz, M. A. Clark, and R. C. Walker, "An extensible interface for QM/MM molecular dynamics simulations with AMBER," J. Comput. Chem. **35**, 95–108 (2014).

[24] P. Sherwood, A. H. de Vries, S. J. Collins, S. P. Greatbanks, N. A. Burton, M. A. Vincent, and I. H. Hillier, "Computer simulation of zeolite structure and reactivity using embedded cluster methods," Faraday Discuss. **106**, 79–92 (1997).

[25] A. H. de Vries, P. Sherwood, S. J. Collins, A. M. Rigby, M. Rigutto, and G. J. Kramer, "Zeolite structure and reactivity by combined quantum-chemical-classical calculations," J. Phys. Chem. B **103**, 6133–6141 (1999).

[26] P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjøvoll, A. Fahmi, A. Schäfer, and C. Lennartz, "QUASI: A general purpose implementation of the QM/MM approach and its application to problems in catalysis," J. Mol. Struct.: THEOCHEM **632**, 1–28 (2003).

[27] P. H. König, M. Hoffmann, T. Frauenheim, and Q. Cui, "A critical evaluation of different QM/MM frontier treatments with SCC-DFTB as the QM method," J. Phys. Chem. B **109**, 9082–9095 (2005).

[28] H. Lin and D. G. Truhlar, "Redistributed charge and dipole schemes for combined quantum mechanical and molecular mechanical calculations," J. Phys. Chem. A **109**, 3991–4004 (2005).

[29] B. Wang and D. G. Truhlar, "Tuned and balanced redistributed charge scheme for combined quantum mechanical and molecular mechanical (QM/MM) methods and fragment methods: Tuning based on the CM5 charge model," J. Chem. Theory Comput. **9**, 1036–1042 (2013).

[30] M. C. R. Melo, R. C. Bernardi, T. Rudack, M. Scheurer, C. Riplinger, J. C. Phillips, J. D. C. Maia, G. B. Rocha, J. V. Ribeiro, J. E. Stone, F. Neese, K. Schulten, and Z. Luthey-Schulten, "NAMD goes quantum: An integrative suite for hybrid simulations," Nat. Methods **15**, 351–354 (2018).

[31] M. Eichinger, P. Tavan, J. Hutter, and M. Parrinello, "A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields," J. Chem. Phys. **110**, 10452–10467 (1999).

[32] P. Amara and M. J. Field, "Evaluation of an *ab initio* quantum mechanical/molecular mechanical hybrid-potential link-atom method," Theor. Chem. Acc. **109**, 43–52 (2003).

[33] G. Koenig, Y. Mei, F. C. Pickard, A. C. Simmonett, B. T. Miller, J. M. Herbert, H. L. Woodcock, B. R. Brooks, and Y. Shao, "Computation of hydration free energies using the multiple environment single system quantum mechanical/molecular mechanical method," J. Chem. Theory Comput. **12**, 332–344 (2016).

[34] J. Chen, Y. Shao, and J. Ho, "Are explicit solvent models more accurate than implicit solvent models? A case study on the Menschutkin reaction," J. Phys. Chem. A **123**, 5580–5589 (2019).

[35] T. J. Giese and D. M. York, "Charge-dependent model for many-body polarization, exchange, and dispersion interactions in hybrid quantum mechanical/molecular mechanical calculations," J. Chem. Phys. **127**, 194101 (2007).

[36] Q. A. Smith, K. Ruedenberg, M. S. Gordon, and L. V. Slipchenko, "The dispersion interaction between quantum mechanics and effective fragment potential molecules," J. Chem. Phys. **136**, 244107 (2012).

[37] E. R. Kuechler, T. J. Giese, and D. M. York, "Charge-dependent many-body exchange and dispersion interactions in combined QM/MM simulations," J. Chem. Phys. **143**, 234111 (2015).

[38] L. V. Slipchenko, M. S. Gordon, and K. Ruedenberg, "Dispersion interactions in QM/EFP," J. Phys. Chem. A **121**, 9495–9507 (2017).

[39] T. Giovannini, P. Lafiosca, and C. Cappelli, "A general route to include Pauli repulsion and quantum dispersion effects in QM/MM approaches," J. Chem. Theory Comput. **13**, 4854–4870 (2017).

[40] C. Curutchet, L. Cupellini, J. Kongsted, S. Corni, L. Frediani, A. H. Steindal, C. A. Guido, G. Scalmani, and B. Mennucci, "Density-dependent formulation of dispersion–repulsion interactions in hybrid multiscale quantum/molecular mechanics (QM/MM) models," J. Chem. Theory Comput. **14**, 1671–1681 (2018).

[41]H. Gökcan, E. Kratz, T. A. Darden, J.-P. Piquemal, and G. A. Cisneros, "QM/MM simulations with the Gaussian electrostatic model: A density-based polarizable potential," J. Phys. Chem. Lett. **9**, 3062–3067 (2018).
[42]C. I. Viquez Rojas, J. Fine, and L. V. Slipchenko, "Exchange-repulsion energy in QM/EFP," J. Chem. Phys. **149**, 094103 (2018).
[43]T. J. Giese and D. M. York, "Ambient-potential composite Ewald method for *ab initio* quantum mechanical/molecular mechanical molecular dynamics simulation," J. Chem. Theory Comput. **12**, 2611–2632 (2016).
[44]N. Ferré and J. G. Ángyán, "Approximate electrostatic interaction operator for QM/MM calculations," Chem. Phys. Lett. **356**, 331–339 (2002).
[45]D. Riccardi, P. Schaefer, and Q. Cui, "pK_a calculations in solution and proteins with QM/MM free energy perturbation simulations: A quantitative test of QM/MM protocols," J. Phys. Chem. A **109**, 17715–17733 (2005).
[46]G. Hou, X. Zhu, M. Elstner, and Q. Cui, "A modified QM/MM Hamiltonian with the self-consistent-charge density-functional-tight-binding theory for highly charged QM regions," J. Chem. Theory Comput. **8**, 4293–4304 (2012).
[47]A. Laio, J. VandeVondele, and U. Rothlisberger, "A Hamiltonian electrostatic coupling scheme for hybrid Car–Parrinello molecular dynamics simulations," J. Chem. Phys. **116**, 6941–6947 (2002).
[48]K. Nam, J. Gao, and D. M. York, "An efficient linear-scaling Ewald method for long-range electrostatic interactions in combined QM/MM calculations," J. Chem. Theory Comput. **1**, 2–13 (2005).
[49]R. C. Walker, M. F. Crowley, and D. A. Case, "The implementation of a fast and accurate QM/MM potential method in amber," J. Comput. Chem. **29**, 1019–1031 (2008).
[50]Z. C. Holden, R. M. Richard, and J. M. Herbert, "Periodic boundary conditions for QM/MM calculations: Ewald summation for extended Gaussian basis sets," J. Chem. Phys. **139**, 244108 (2013).
[51]Z. C. Holden, B. Rana, and J. M. Herbert, "Analytic gradient for the QM/MM-Ewald method using charges derived from the electrostatic potential: Theory, implementation, and application to *ab initio* molecular dynamics simulation of the aqueous electron," J. Chem. Phys. **150**, 144115 (2019).
[52]Y. Zhou, S. Wang, Y. Li, and Y. Zhang, "Born–Oppenheimer *ab initio* QM/MM molecular dynamics simulations of enzyme reactions," in *Methods Enzymol* (Elsevier, 2016), Vol. 577, pp. 105–118.
[53]D. Fang, R. E. Duke, and G. A. Cisneros, "A new smoothing function to introduce long-range electrostatic effects in QM/MM calculations," J. Chem. Phys. **143**, 044103 (2015).
[54]E. G. Kratz, R. E. Duke, and G. A. Cisneros, "Long-range electrostatic corrections in multipolar/polarizable QM/MM simulations," Theor. Chem. Acc. **135**, 166 (2016).
[55]T. Vasilevskaya and W. Thiel, "Periodic boundary conditions in QM/MM calculations: Implementation and tests," J. Chem. Theory Comput. **12**, 3561–3570 (2016).
[56]X. Pan, E. Rosta, and Y. Shao, "Representation of the QM subsystem for long-range electrostatic interaction in non-periodic *ab initio* QM/MM calculations," Molecules **23**, 2500 (2018).
[57]M. J. Field, P. A. Bash, and M. Karplus, "A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations," J. Comput. Chem. **11**, 700–733 (1990).
[58]C. F. Sanz-Navarro, R. Grima, A. García, E. A. Bea, A. Soba, J. M. Cela, and P. Ordejón, "An efficient implementation of a QM-MM method in SIESTA," Theor. Chem. Acc. **128**, 825–833 (2011).
[59]T. Okamoto, K. Yamada, Y. Koyano, T. Asada, N. Koga, and M. Nagaoka, "A minimal implementation of the AMBER-Gaussian interface for *ab initio* QM/MM-MD simulation," J. Comput. Chem. **32**, 932–942 (2011).
[60]N. M. Thellamurege, D. Si, F. Cui, H. Zhu, R. Lai, and H. Li, "QuanPol: A full spectrum and seamless QM/MM program," J. Comput. Chem. **34**, 2816–2833 (2013).
[61]S. Riahi and C. N. Rowley, "The CHARMM-TURBOMOLE interface for efficient and accurate QM/MM molecular dynamics, free energies, and excited state properties," J. Comput. Chem. **35**, 2076–2086 (2014).
[62]K. Nam, "Acceleration of *ab initio* QM/MM calculations under periodic boundary conditions by multiscale and multiple time step approaches," J. Chem. Theory Comput. **10**, 4175–4183 (2014).
[63]Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. W. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, "Advances in molecular quantum chemistry contained in the Q-CHEM 4 program package," Mol. Phys. **113**, 184–215 (2015).
[64]T. Kubař, K. Welke, and G. Groenhof, "New QM/MM implementation of the DFTB3 method in the GROMACS package," J. Comput. Chem. **36**, 1978–1989 (2015).
[65]E. G. Kratz, A. R. Walker, L. Lagardère, F. Lipparini, J.-P. Piquemal, and G. Andrés Cisneros, "LICHEM: A QM/MM program for simulations with multipolar and polarizable force fields," J. Comput. Chem. **37**, 1019–1029 (2016).
[66]P. Ojeda-May and K. Nam, "Acceleration of semiempirical QM/MM methods through message passage interface (MPI), hybrid MPI/open multiprocessing, and self-consistent field accelerator implementations," J. Chem. Theory Comput. **13**, 3525–3536 (2017).
[67]D. Loco, L. Lagardere, G. A. Cisneros, G. Scalmani, M. Frisch, F. Lipparini, B. Mennucci, and J.-P. Piquemal, "Towards large scale hybrid QM/MM dynamics of complex systems with advanced point dipole polarizable embeddings," Chem. Sci. **10**, 7200–7211 (2019).
[68]A. de la Lande, A. Alvarez-Ibarra, K. Hasnaoui, F. Cailliez, X. Wu, T. Mineva, J. Cuny, P. Calaminici, L. López-Sosa, G. Geudtner, I. Navizet, C. Garcia Iriepa, D. R. Salahub, and A. M. Köster, "Molecular simulations with in-deMon2k QM/MM, a tutorial-review," Molecules **24**, 1653 (2019).
[69]G. A. Cisneros, M. Karttunen, P. Ren, and C. Sagui, "Classical electrostatics for biomolecular simulations," Chem. Rev. **114**, 779–814 (2014).
[70]Y. Kawashima, K. Ishimura, and M. Shiga, "*Ab initio* quantum mechanics/molecular mechanics method with periodic boundaries employing Ewald summation technique to electron-charge interaction: Treatment of the surface-dipole term," J. Chem. Phys. **150**, 124103 (2019).
[71]F. Melaccio, M. Olivucci, R. Lindh, and N. Ferré, "Unique QM/MM potential energy surface exploration using microiterations: Exploration of QM/MM energy surface," Int. J. Quantum Chem. **111**, 3339–3346 (2011).
[72]F. Aquilante, J. Autschbach, R. K. Carlson, L. F. Chibotaru, M. G. Delcey, L. De Vico, I. Fdez. Galván, N. Ferré, L. M. Frutos, L. Gagliardi, M. Garavelli, A. Giussani, C. E. Hoyer, G. Li Manni, H. Lischka, D. Ma, P. Å. Malmqvist, T. Müller, A. Nenov, M. Olivucci, T. B. Pedersen, D. Peng, F. Plasser, B. Pritchard, M. Reiher, I. Rivalta, I. Schapiro, J. Segarra-Martí, M. Stenrup, D. G. Truhlar, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov, V. P. Vysotskiy, O. Weingart, F. Zapata, and R. Lindh, "Molcas 8: New capabilities for multiconfigurational quantum chemical calculations across the periodic table," J. Comput. Chem. **37**, 506–541 (2016).
[73]F. Aquilante, J. Autschbach, A. Baiardi, S. Battaglia, V. A. Borin, L. F. Chibotaru, I. Conti, L. De Vico, M. Delcey, I. F. Galván, N. Ferré, L. Freitag, M. Garavelli,

X. Gong, S. Knecht, E. D. Larsson, R. Lindh, M. Lundberg, P. Å. Malmqvist, A. Nenov, J. Norell, M. Odelius, M. Olivucci, T. B. Pedersen, L. Pedraza-González, Q. M. Phung, K. Pierloot, M. Reiher, I. Schapiro, J. Segarra-Martí, F. Segatta, L. Seijo, S. Sen, D.-C. Sergentu, C. J. Stein, L. Ungur, M. Vacher, A. Valentini, and V. Veryazov, "Modern quantum chemistry with [Open]Molcas," J. Chem. Phys. **152**, 214117 (2020).

[74] J. A. Rackers, Z. Wang, C. Lu, M. L. Laury, L. Lagardère, M. J. Schnieders, J.-P. Piquemal, P. Ren, and J. W. Ponder, "Tinker 8: Software tools for molecular design," J. Chem. Theory Comput. **14**, 5273–5289 (2018).

[75] K. Schwinn, N. Ferré, and M. Huix-Rotllant, "Analytic QM/MM atomic charge derivatives avoiding the scaling of coupled perturbed equations with the MM subsystem size," J. Chem. Phys. **151**, 041102 (2019).

[76] K. Schwinn, N. Ferré, and M. Huix-Rotllant, "Efficient analytic second derivative of electrostatic embedding QM/MM energy: Normal mode analysis of plant cryptochrome," J. Chem. Theory Comput. **16**, 3816–3824 (2020).

[77] K. Schwinn, N. Ferré, and M. Huix-Rotllant, "UV-visible absorption spectrum of FAD and its reduced forms embedded in a cryptochrome protein," Phys. Chem. Chem. Phys. **22**(22), 12447–12455 (2020).

[78] C. M. Breneman and K. B. Wiberg, "Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis," J. Comput. Chem. **11**, 361–373 (1990).

[79] U. C. Singh and P. A. Kollman, "An approach to computing electrostatic charges for molecules," J. Comput. Chem. **5**, 129–145 (1984).

[80] B. H. Besler, K. M. Merz, and P. A. Kollman, "Atomic charges derived from semiempirical methods," J. Comput. Chem. **11**, 431–439 (1990).

[81] W. D. Cornell, P. Cieplak, C. I. Bayly, and P. A. Kollman, "Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation," J. Am. Chem. Soc. **115**, 9620–9631 (1993).

[82] D. Beglov and B. Roux, "Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations," J. Chem. Phys. **100**, 9050–9063 (1994).

[83] A. Aleksandrov and M. Field, "Efficient solvent boundary potential for hybrid potential simulations," Phys. Chem. Chem. Phys. **13**(22), 10503 (2011).

[84] W. Im, S. Bernèche, and B. Roux, "Generalized solvent boundary potential for computer simulations," J. Chem. Phys. **114**, 2924–2937 (2001).

[85] P. Schaefer, D. Riccardi, and Q. Cui, "Reliable treatment of electrostatics in combined QM/MM simulation of macromolecules," J. Chem. Phys. **123**, 014905 (2005).

[86] T. Benighaus and W. Thiel, "Efficiency and accuracy of the generalized solvent boundary potential for hybrid QM/MM simulations: Implementation for semiempirical Hamiltonians," J. Chem. Theory Comput. **4**, 1600–1609 (2008).

[87] T. Benighaus and W. Thiel, "A general boundary potential for hybrid QM/MM simulations of solvated biomolecular systems," J. Chem. Theory Comput. **5**, 3114–3128 (2009).

[88] T. Benighaus and W. Thiel, "Long-range electrostatic effects in QM/MM studies of enzymatic reactions: Application of the solvated macromolecule boundary potential," J. Chem. Theory Comput. **7**, 238–249 (2011).

[89] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, and M. Parrinello, "PLUMED: A portable plugin for free-energy calculations with molecular dynamics," Comput. Phys. Commun. **180**, 1961–1972 (2009).

[90] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," SIAM J. Numer. Anal. **10**, 413–432 (1973).

[91] P. J. Steinbach and B. R. Brooks, "New spherical-cutoff methods for long-range forces in macromolecular simulation," J. Comput. Chem. **15**, 667–683 (1994).

[92] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, "The AMBER biomolecular simulation programs," J. Comput. Chem. **26**, 1668–1688 (2005).

[93] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," Phys. Rev. A **38**, 3098–3100 (1988).

[94] A. D. Becke, "A new mixing of Hartree–Fock and local density-functional theories," J. Chem. Phys. **98**, 1372–1377 (1993).

[95] C. Lee, W. Yang, and R. G. Parr, "Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density," Phys. Rev. B **37**, 785–789 (1988).

[96] W. J. Hehre, R. Ditchfield, and J. A. Pople, "Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules," J. Chem. Phys. **56**, 2257–2261 (1972).

[97] P. C. Hariharan and J. A. Pople, "The influence of polarization functions on molecular orbital hydrogenation energies," Theor. Chim. Acta **28**(3), 213–222 (1973).

[98] R. A. Kendall, T. H. Dunning, and R. J. Harrison, "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions," J. Chem. Phys. **96**, 6796–6806 (1992).

[99] R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig, and A. D. MacKerell, "Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ_1 and χ_2 dihedral angles," J. Chem. Theory Comput. **8**, 3257–3273 (2012).

[100] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. Mackerell, "CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields," J. Comput. Chem. **31**, 671 (2009).

[101] K. Vanommeslaeghe and A. D. MacKerell, "Automation of the CHARMM general force field (CGenFF) I: Bond perception and atom typing," J. Chem. Inf. Model. **52**, 3144–3154 (2012).

[102] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," J. Chem. Phys. **79**, 926–935 (1983).

[103] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, and C. Simmerling, "ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB," J. Chem. Theory Comput. **11**, 3696–3713 (2015).

[104] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case, "Development and testing of a general AMBER force field," J. Comput. Chem. **25**, 1157–1174 (2004).

[105] G. M. Torrie and J. P. Valleau, "Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid," Chem. Phys. Lett. **28**, 578–581 (1974).

[106] M. R. Shirts and J. D. Chodera, "Statistically optimal analysis of samples from multiple equilibrium states," J. Chem. Phys. **129**, 124105 (2008).

[107] X. Pan, P. Li, J. Ho, J. Pu, Y. Mei, and Y. Shao, "Accelerated computation of free energy profile at *ab initio* quantum mechanical/molecular mechanical accuracy via a semi-empirical reference potential. II. Recalibrating semi-empirical parameters with force matching," Phys. Chem. Chem. Phys. **21**(37), 20595–20605 (2019).

[108] U. Ryde, "How many conformations need to be sampled to obtain converged QM/MM energies? The curse of exponential averaging," J. Chem. Theory Comput. **13**, 5745–5752 (2017).

[109] M. L. Sanchez, M. E. Martin, M. A. Aguilar, and F. J. Olivares del Valle, "Solvent effects by means of averaged solvent electrostatic potentials: Coupled method," J. Comput. Chem. **21**(9), 705–715 (2000).

[110] E. Rosta, M. Haranczyk, Z. T. Chu, and A. Warshel, "Accelerating QM/MM free energy calculations: Representing the surroundings by an updated mean charge distribution," J. Phys. Chem. B **112**, 5680–5692 (2008).

[111] Y. Orozco-Gonzalez, M. Manathunga, M. D. C. Marín, D. Agathangelou, K.-H. Jung, F. Melaccio, N. Ferré, S. Haacke, K. Coutinho, S. Canuto, and M. Olivucci, "An average solvent electrostatic configuration protocol for QM/MM free energy optimization: Implementation and application to rhodopsin systems," J. Chem. Theory Comput. **13**, 6391–6404 (2017).

[112] D. Nocito and G. J. O. Beran, "Averaged condensed phase model for simulating molecules in complex environments," J. Chem. Theory Comput. **13**, 1117–1129 (2017).

[113] N. Okuyama-Yoshida, K. Kataoka, M. Nagaoka, and T. Yamabe, "Structure optimization via free energy gradient method: Application to glycine zwitterion in aqueous solution," J. Chem. Phys. **113**, 3519–3524 (2000).

[114]H. Hirao, Y. Nagae, and M. Nagaoka, "Transition-state optimization by the free energy gradient method: Application to aqueous-phase Menshutkin reaction between ammonia and methyl chloride," Chem. Phys. Lett. **348**, 350–356 (2001).

[115]H. C. Georg and S. Canuto, "Electronic properties of water in liquid environment. A sequential QM/MM study using the free energy gradient method," J. Phys. Chem. B **116**, 11247–11254 (2012).